

## 音声認識技術を応用した発音練習教材の 開発と設計

—留学生のための日本語の音声指導に向けて—

高木 南欧子

In order to conduct Japanese phonetics education with international students as the subject, this study examined development of study contents for practicing prosody that are suited to undergraduate international students. For this purpose, I developed prosody practicing aids in applying phonetic recognition techniques. For the CALL System, I utilized the pre-existing “Japanese CALL System”. As a result, I found that the expressions available were insufficient for use on campus, each sentence was comparatively short, and there were cases where pronunciation was not evaluated precisely by the CALL System. I therefore included three situations that would be familiar for international students, and created 40 sentences of a maximum of 40 morae (syllables) . Upon inserting them into the CALL System, I analyzed the problems associated with pronunciation of the added sentences and applied settings that reflected pronunciation evaluation. Also, in order for adjustments to be possible in actual use, I applied three settings for the level of strictness used as criteria for the automatic evaluation. After implementation of the study contents in class, I found that determining what problems faced each student became clearer, and there were improvements in how students could practice individually, but, on the other hand, there were problems with the system recognitionising differences in voiced and voiceless sounds as a sentence became longer.

キーワード：発音練習, 音声認識技術, アクセント, イントネーション,  
CALL

## 1. はじめに

大学受験を目標とした日本語教育において、言語活動「読む」、「書く」、「聞く」、「話す」の4領域のうち、「話す」練習は優先順位が低くなる傾向がある。しかしながら大学入学後は、ゼミナール等における研究活動、インターンシップへの参加、就職活動などにおいて、抽象的な概念の説明や裏付けのある討論への参加を求められることがある。このような難易度の高いレベルにおける発話を行うためには、少なくとも「外国人の日本語に慣れていない人にもわかるが、母語の影響が残っている」「ときどきつかえることはあるが、一人でどんどん話せる」という上級<sup>i</sup>の能力が必要である。しかし、学部留学生の入学時の日本語発話能力は、「外国人の日本語に慣れている人にはわかる」「つかえることが多いし、一人で話しつつづけることは難しい」という中級<sup>ii</sup>のレベルにあることが多い。しかし、日本語を専攻していない留学生の場合、学年が上がると外国語としての日本語は履修の必要性がなくなることが多い。3、4年次におけるゼミナール活動や就職活動などに取り組むようになると、口頭能力の不足が顕在化することがあるが、その時点ではすでに学習するための時間的余裕はなくなっている。また、この問題をサポートする仕組みがカリキュラムとして存在しないケースも多い。口頭運用能力の問題は、留学生を対象とした日本語教育をとりまく課題の一つとなっている。

## 2. 留学生の卒業後の進路

日本語非母語話者の日本語は、レベルが上がるにしたがって厳しく評価されるという指摘がある（小河原 2001a, 2001b）。ビジネスなどの仕事の場面において、日本語母語話者による評価がより厳しくなるのであれば、留学生の卒業後の進路も考慮に入れた音声教育を在学中から行う必要性がある。アジア人材資金構想プロジェクトサポートセンター（2011）においても、企業の採用にあたっては、「聞く力」、「話す力」が留学生に求められること、就業後は技術職を除き、多くの場合、日本語能力試験1級以上の日本語能力が求められる傾向にあることが報告されている。

日本学生支援機構は、「外国人留学生進路状況・学位授与状況調査」を公表しており、卒業後の留学生の進路状況を見ることができる。調査結果が発表されている直近5年間の2010年から2014年を見ると、大学卒業後に日本国内、または出身国（地域）で就職する割合は図1、図2のように

なっている。なお、図1、図2は卒業後の就職先の国を在学段階別に見ているものである。本稿では、大学の学部4年間での学びの結果とその進路状況を見るため、扱う在学の種類は「学部」、「修士」、「博士」の3つに限定し、抽出した。これを見ると、博士課程を修了した留学生は、出身国（地域）に帰国して就職する割合が高く、修士課程を修了した留学生は日本国内で就職する割合がやや多いことが分かる。学部留学生に至っては、卒業後、出身国（地域）で就職するよりも、日本国内で就職する割合の方が多くなっている。

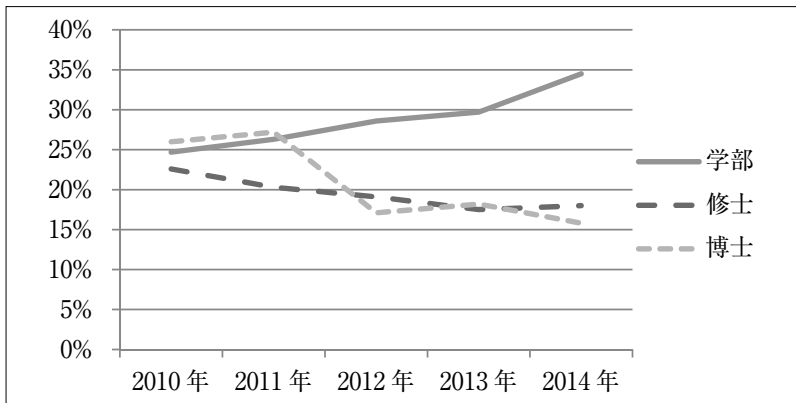


図1 卒業後、日本国内で就職する留学生の割合<sup>iii</sup>  
 (日本学生支援機構「外国人留学生進路状況・学位授与状況調査」平成22年～平成26年の数値を在学段階別にまとめたもの),

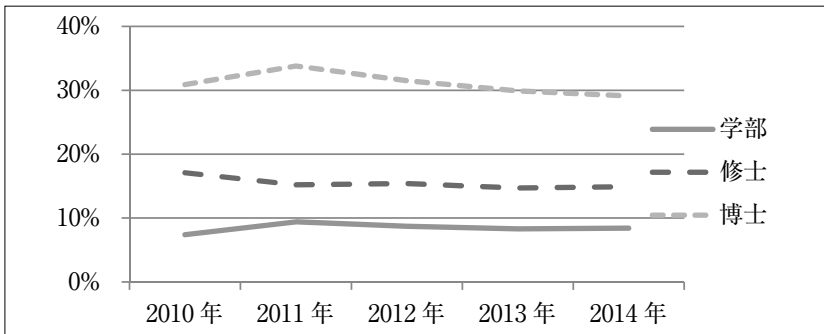


図2 卒業後、出身国（地域）に帰国して就職する留学生の割合

図3は在学段階を学部限定し、学部留学生の主な進路先がどのような割合で推移しているかを示したものである。これを見ると、「日本国内で就職」の数値は上昇傾向が見られるものの、平均すると30%前後であり、進路先としては一番高い割合であることが分かる。「日本国内で就職」,「日本国内で進学」を合わせると、学部留学生の約40%あまりが卒業後、日本国内に留まることを選択していることになる。仕事や高度な研究活動の場において、日本語がコミュニケーションツールとして長く留学生の中に位置づけられていくのであれば、在学中における音声教育は、より重きを置かれてしかるべきだと考えられる。

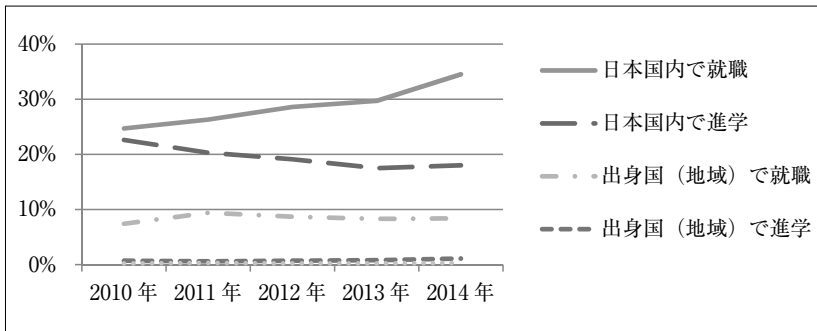


図3 学部卒業後の留学生の進路の割合<sup>iv</sup>

### 3. 音声教育について

発話能力を上げる方法には、聴解練習やCALLの利用による発話練習、シャドーイングなどがあり、これらはすでに様々な実践がなされている。特にシャドーイングは、単音の間違いの減少 (萩原 2005)、イントネーションの改善 (高橋・松崎 2007) などにおいて効果があることが報告されている。しかしながら、阿・林 (2010) では文末イントネーションの学習効果が中国語母語話者にはあまり見られなかった事例が報告されており、母語によって効果が異なる可能性が指摘されている。高橋・松崎 (2007) は、シャドーイングの問題点として、シャドーイングは学習者の聞き取り能力を前提としているため、自律学習としては限界があることを同時に指摘している。シャドーイングは、聞こえてくる音声を即時的に再生していく練習方法であるため、聴解の能力が高くない学習者の場合、高い負荷により

学習者自身の発話のモニタリングが困難になることが原因の一つと考えられる。

学習に対する学習者自身のモニタリングの働きについては、様々な分野でも研究がされているが、音声教育においても、学習者自身のモニタリングを重視すべきだとする考え方がある。小河原（2009）は「自己モニター」を、学習者自身が自己のパフォーマンスの問題点を認識し、模索しながら適切な自己評価を通して発音を自己修正することとしている。このような自己モニターを活用した音声教育に、河野（2015,2016）の単音、名詞のアクセントについての実践報告がある。これは学習者が自身の発音の問題点を認識し、発音を修正していくという方法である。しかし、自己モニターを機能させるには、自己モニターとなる基準を個々の学習者がそれぞれ独自に確立させていくことが前提となる。

自己モニターの基準を作る過程においては、学習者が個々に発音の試行錯誤を行い、他者などからの客観的な評価を得ることが必要である。しかし、学習者によっては他者からの発音評価を得ることに心理的な負担を感じる場合がある。太田・工藤（2012）では、外国人向け日本語発音矯正システム「AmiVoice CALL Web-Japanese」（以下、AmiVoice）を用いた実践報告がなされている。対象は、日本語能力試験N2程度の学習をして入学してきた新入留学生である。この日本語CALLシステムでは、音声入力後、機械により自動的に発音評価が表示されるようになっている。そのため、学習者自身が感じる対人的な心理的負担は軽減される。調査結果では、インターフェイスや項目リストの使いにくさが問題点として上げられていたものの、一定の成果があげられたことが報告されている。太田・工藤（2012）における対象学習者の層や環境は、本稿で目指す留学生の日本語教育に共通する点がある。また、教材の性質として、自学学習がある程度可能であるということも本研究が目指す音声教材と一致する。このことから、AmiVoiceを利用した発音教材コンテンツの開発を目指し、開発計画を進める上での検討を行うこととした。

### 3. 音声教材としての利用

#### 3.1 音声認識技術と単音の評価

「音声認識は、観測された音声波形から、発声者が何を話した（話そうとした）のかを推定する」（古井2000）ものである。音声認識技術は近年、

書き起こしなどを目的とした音声入力などで用いられている技術であり、現在も急速な技術開発が進められている。ここでは、自動音声書き起こしソフトとしての AmiVoice エンジンを見ることにより、日本語 CALL システムの音声認識の仕組みを見ることとする。

図4は、自動音声書き起こしソフトとしての AmiVoice エンジンのイメージである<sup>v</sup>。音声入力では、まず発声者がマイクロホンに向かって発声を行う。発声された音声は音声波形に変換される。波形は一連の音素の連続体であるが、この波形は音響モデルにおいて音素単位で切り取られる。そして、周波数成分などの分析から各音素の特徴量が抽出され、発声された音が何であったかが判別される<sup>vi</sup>。

日本語 CALL システムとしての AmiVoice は、この音声認識技術を利用したもので、さらに日本語学習者誤用データベースを組み込んだ（岩田・清水・西村 2009）ことによって発音評価を可能としている。音声入力ソフトと異なる点は、音声入力は自由な発話によるものではなく、あらかじめコンテンツに搭載された読み上げ文に限定されるという点である。学習者が読み上げた音声（音素）と、読み上げ文を構成する音素が比較され、その近さによって評価される。この評価は、音声入力後、自動的に行われ、学習者に表示される。



図4 自動音声入力ソフトとしての AmiVoice

音声評価は、入力した音声の音素と、読み上げ文が想定している音素モデルとの近さによって評価されるが、近さは遠さと表現することもできる。人間は個々に発声器官を持つが、それぞれ少々の特徴を持つ。また、発音

時の癖や特徴、発音時の調音点や調音法の微妙な違い、環境などからも影響を受け、音素の特徴量の差が生まれる。そのような中で生まれた多様な音を、人間はある程度の幅を持って捉え、音の同定を無意識に行う。音声認識においても、一つの音素につき、ある程度の幅を持たせた範囲内のものを同じ音素と判定させる。しかしながら、入力された音素と音素モデルの特徴量が数値的には近く、自動判定では同じ音と判別されたとしても、人間の耳で聞くと違和感を強く感じられることもある。現時点における音響エンジンでは性能等の限界があるため、人間の耳で間違いと判定されるもの全てが、機械で同じように間違いと判定できるわけではない。

AmiVoiceでは、「発音チェック」練習メニューにおいて、マイクロホンを通して入力した音声の発音評価(図5①)が自動的に表示される。入力音声のうちエラーがあると判定された場合、エラーの箇所が表示される(図5⑤)。表示されるエラーの種類は、38種類ある。



図5 発音チェックの画面<sup>vii</sup>

エラーの種類は、主に子音に関わるもの、母音に関わるもの、調音方法に関わるものに分類できる。日本語学習者の発音によく見られるような「つ」・「ちゅ」、有声音・無声音、母音の長短、促音の有無等の指摘が可能(岩田・清水・西村 2009)である。しかしながら、例えば「な行」を「ら行」で発音している場合のエラーの指摘は可能であるが、反対に「ら行」を「な行」で発音している場合のエラーは性能による限界から指摘できないなど、

日本語学習者によくある発音の間違いであっても、エラーとして処理できない音素もあるため、使用上の制約がある。

## 4.2 韻律の評価

佐藤（1995）は、韓国語話者、中国語話者、東京方言話者の各1名ずつをインフォーマントとし、発話音声日本語母語話者に与える不自然さの印象について調査を行っている。そこで、韓国語話者、中国語話者ともに、日本語の不自然さに影響を与えるのは単音よりも韻律の方が強いこと、高さ・長さ・強さの3つの韻律的要素のうち、高さが評価に与える影響が最も大きいことを明らかにしている。

AmiVoiceでは、単音の評価は可能であるが、アクセントやイントネーションといった韻律の評価システムはない。そのため、図5の②プロソディグラフ<sup>viii</sup>、③モデル音声と入力音声の波形の比較、④モデル音声と入力音声のピッチ曲線の比較が韻律を確認する手がかりとなる。松崎・高橋（2015）は、AmiVoiceの音声認識エンジンを組み込み、入力音声の韻律が自動的に表示される韻律学習CALL教材であるProsody Tuner<sup>ix</sup>を開発し、図5にはない様々な機能を加えている。その機能は韻律に関する自己モニターを促し、実際に聞こえる聴覚印象に学習が発話を近づけるための方策であるが、その一つに「おそ下がり」に対応したプロソディグラフがある。「おそ下がり」は、物理的なピッチ下降が聴覚印象より1拍程度遅れて生じる聴覚現象である。この現象を知らずにピッチ曲線だけを見ると、モデル音声の韻律だけでなく、学習者が自身の韻律に対し、間違った理解をする恐れがある。同時に、誤った自己モニター基準の形成につながることも意味する。

しかしながら、Prosody Tunerの利用は、本研究におけるCALL教室の環境下ではいくつか難しい点がある。そのため、本研究での韻律学習においては、図5における「発音チェック」のみで行い、「おそ下がり」等の対応については運用段階で検討することとした。

## 5. 試行実施

追加コンテンツの開発にあたっては、既存のAmiVoiceにおける発音学習の効果と課題をあぶり出す必要があった。そのため、中・上級の留学生6名を対象とした「音声」のクラスにおいて、シラバスの一部に既存の



AmiVoice（試用版）を利用した学習を組み込み、発音教育を奥村（2016）が行った。この試行的実施は筆者自身が行ったものではないが、本開発計画の一部であり、基礎的資料となるため、ここで簡単に紹介しておく。試行実施後のアンケートにおいて、学習者からは①個人の発音上の課題に気づき、より高い目標を持つようになった（5名）、②学習目標への意欲が高くなった（5名）、③苦手な発音に向上の兆しが感じられるようになった（4名）という評価を得た。①の発音上の課題に対する気づきという項目では、同じエラーが何度も表示されることにより、苦手とする発音が明確になったという回答が得られた。②の学習意欲という項目では、日本語 CALL における練習時間だけでなく、教室内での他の活動や、教室外での発音・ヒアリングにも気を付けるようになったという回答があった。③苦手な発音の向上という項目は、AmiVoice における「発音チェック」の使用に、教室活動において自己モニターの意識化を組み合わせさせた結果である。調音点、調音法についての方法を意識する活動を設けた後、再度「発音チェック」メニューにおいて練習を行った結果、評価の向上が見られた。顕著な例では、「で」と「れ」の間違いにおいて、1回目の正答率が60%、2回目の正答率が50%、3回目の正答率が80%と推移したケースも見られた。しかしながら、回数を重ねても評価がなかなか上がらない学習者がある、良い評価が出すぎる（評価が甘い）ことがある、中・上級レベルとしては一つのセンテンスが短く、大学生活で使用するものが少ない、という課題があることも明らかになった。

そこで、この試行結果における課題を踏まえ、本発音教材の開発においては、次の3点について配慮を行うこととした。まず1点目は、評価と指導の対象とする項目と許容範囲の設定である。母語や個人的な癖等により、発音の改善が非常に困難であるというケースは、しばしば見られることがある。音素によっては、語頭あるいは語中など、位置によってエラーの出現率が異なるものがある。そのような音素の場合、エラーの出現率が低い語彙で練習をした後、エラーの出現率の高い語彙などで練習を行うなどの方法もある。しかし、これらは改善にはある程度の時間を要する傾向がある。また、発音を意識しすぎるが故に心理的な負担となり、悪影響を与えることもある。コミュニケーションという視点から考えると、意思伝達に大きな影響を与えないレベルであるならば、発音時の音素のずれを個性として捉えようとする考え方が一般的である。

次に、個々の教室活動をコース全体として俯瞰し、各科目、セメスター間における順次性と学習の意識化への配慮を行うこととした。コースの開始時点においては、発音の改善が難しい場合もあることを含め、発音そのものに関する意識と問題点を前提として教員も学習者も共有が必要である。また、学習者に心理的な負担がかかりすぎないように、シラバス、コースデザイン全体の中で長期的な配慮を行うことも重要であろう。音声の授業などにおいては、口腔内の構造の意識化、聞き分け練習、スピーチやシャドーイングなどシラバスを整備しつつ、目標の設定を段階的に行っていくこととした。

最後に、良い評価が出すぎることへの対応である。この問題への対応としては、実装の際に音声評価の難易度を3段階設定することとした。また、その他として、ある程度の長さがあるセンテンス、文型として難易度の高いセンテンスが不足していることもうかがうことができたため、最長40拍の長さの文を設定し、留学生にとって必要性が高い、あるいは利用頻度が高い場面を想定した内容を想定した。これら3つの配慮を条件を開発コンテンツに反映することとした。

## 6. 追加コンテンツの作成の方針

「5. 試行実施」で見た留学生のニーズに応えるため、中・上級レベルの留学生を対象としたコンテンツを40センテンス追加することとした。追加コンテンツは、日本に中・長期に渡り滞在する留学生の学生生活を支援するためのものとし、日本語科目の履修が終わった後も自律学習が可能で、かつ自己モニターの形成につながることを目指す教材とした。また、開発にあたっては、教員から見た教育上の望ましさより、大学生であれば何をどう話すか、という学生からの視点を重視し、開発の方針を立てることとした。ここで立てた方針は次のとおりである。(1) 中・上級日本語レベルにある留学生を対象する、(2) 自律学習としての利用も可能とする、(3) 留学生にとって身近な場面<sup>x</sup>における読み上げ文とする、(4) モデル音声の発話者は、首都圏出身の留学生と日本語母語話者である大学生(20代前半の男女)とする、(5) アクセント、イントネーションは東京方言とする、(6) おそ下がりが生じる読み上げ文においては、聴覚印象に依拠したプロソディグラフを作成する、(7) 入力音声の発音評価を試行版より厳しくする、(8) 平板化などアクセントの揺れや変化が見られる語彙に関しては、

日本語を母語とする大学生が一般に使用する型を用いる、(9) 文末にむかっ  
ての音量の減退は、(8)と同様に、大学生が一般にインタラクションの中  
で発話するような自然なイントネーションにする、(10) センテンスの長  
さは表示の限度上 40 拍とする、(11) 留学生が苦手とする上昇調質問文を  
入れる。

## 7. 発音評価のための調査

### 7.1 不自然な発音の調査

自動的に音声の評定を行うためには、あらかじめ想定される発音のエ  
ラー情報を各読み上げ文に情報として付与しておく必要がある。「4-1. 音  
声認識技術と単音の評価」で見たように、この設定がないと、極端に言えば、  
入力音声の音素の特徴量がモデル音素の特徴量と完全に一致しない限りは  
高い評価がでない、あるいは、厳密でない発音でも高い評価が出る可能  
性が高くなる。この設定を行うため、留学生の発音を録音し、不自然に聞こ  
える音素の抽出を行った。この結果の分類は、読み上げ文に付与するた  
めに使用するが、後々、教室における音声指導においても活用する。

### 7.2 調査と方法

読み上げ文に付与する発音の誤りのタイプは全部で 38 種類である (巻  
末資料)<sup>xi</sup>。38 種類の内訳は、図 6 のようになっている。38 種類の内、最  
も多いものは子音に関わるもの (例:「ら」を「な」で発音する) で 17 種類、  
次いで拍に関わるもの (「っ」「ん」「ゃゅょ」などの有無によって拍の長  
さが変わる) が 14 種類、母音に関わるもの (う段がお段になる、ないは  
ずの母音が添加されるなど) が 7 種類、有声音・無声音 (例「きく」の「き」  
を有声音で読むなど) 2 種類である。

想定される発音のエラーを抽出するため、留学生 31 名<sup>xii</sup>の協力を得て、  
追加コンテンツの読み上げ文 40 センテンスを読み上げてもらい、読み上  
げ音声を合計で 1240 ファイル収録した。次にこの音声ファイル 1240 個を  
2名の日本語教員が別々に聞き、不自然と思われる発音に情報の抽出を行っ  
た。エラーの抽出の見直しは、全部で 3 回行ったが、留学生の発話傾向から、  
当然エラーが出るであろうと予測された箇所でエラーが感じられない箇所  
もあった。これは、当該音素が比較的安定して発音しやすい位置で出現し  
ていたものであったこと、調査対象者が中・上級レベルで安定した実力が

あったことが要因であると推測された。本研究で目指す音声教材は、中・上級レベルの日本語学習者を対象とするが、初級においても対応を可能とするため、初級レベルで想定されるエラーは補完的に追加を行った。

### 7.3 調査結果

抽出したエラーの数の総数は読み上げ文 40 センテンスに対し、2300 であった<sup>xiii</sup>。抽出したエラーをタイプごとに分類した結果が図 7 である。

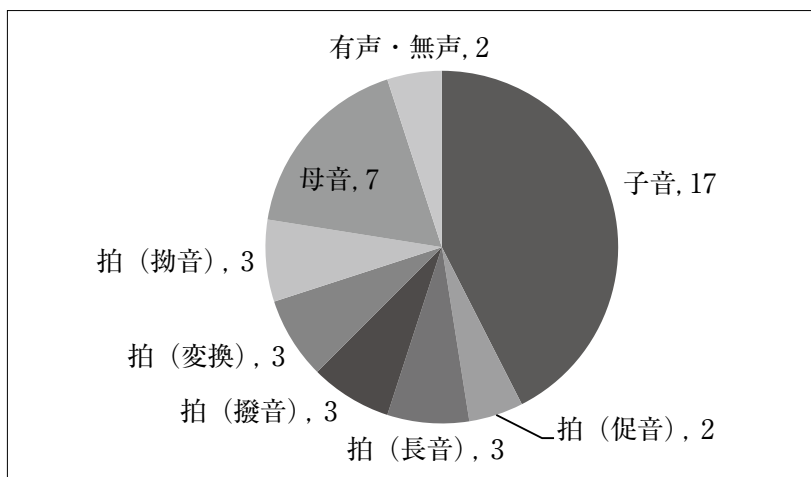


図 6 読み上げ文に付与したエラータイプ 38 種類の内訳

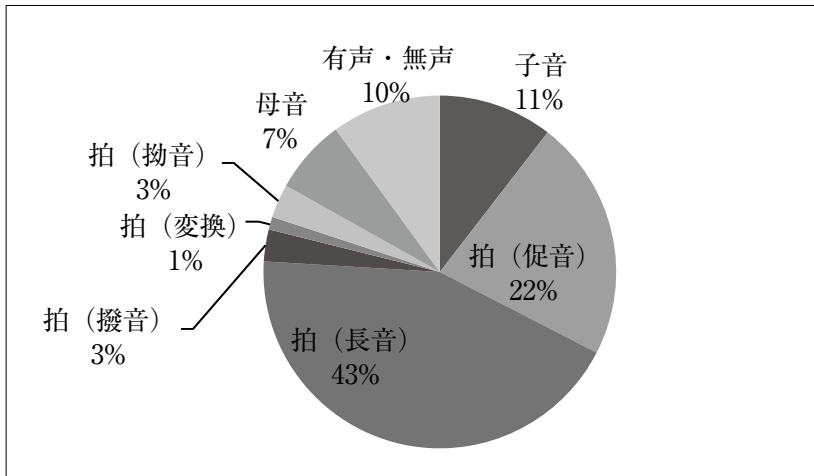


図7 抽出したエラータイプの種類

図7を見ると、最も多いものは拍に関わるもので、全体の72%を占めている。特に拍のエラーのうち、長音に関わるものが43%であり、全体の半分弱となっている。このエラータイプの内訳をさらに調べると、長い拍が短い拍で発音されるタイプのエラーが587、短い拍が長い拍で発音されるタイプのエラーが411となっていた。拍が足りないと感じられるエラーがやや多いと言えよう。また、次に多かったのは促音に関わる拍のエラーであるが、さらに内訳を見ると、「っ」が不足しているタイプのエラーが308、必要のない「っ」が入っているタイプのエラーが200であった。拍の長短と同様、促音に関しても不足していると感じられるエラーの方がやや多くなっている。3番目と4番目はほぼ同数で、子音11%、有声・無声10%であった。子音16種類のうち、最も多かったエラータイプは「だ行」を「ら行」で発音してしまうタイプで59、次いで「し」を[si]で発音してしまうタイプが48であった。子音のエラータイプは16種類、有声・無声のエラータイプは2種類であることを考えると、短期間で発話の自然さの改善を望むのであれば、子音よりも、有声・無声の練習をした方が早いという可能性が考えられよう。

図6は子音に関わるエラーが17種類、拍に関わるエラーが14種類であり、割合にすると子音42%、拍が35%になる。一方、図7では、子音に

関わるエラーが11%、拍に関わるエラー72%と大きく差が開く。音の長さは他の要素に比べ、イントネーションとの関連が強いと考えられる。もし拍の誤りが不自然さを感じさせやすいと言えるのであれば、反対に拍の改善が発音の向上にもつながりやすいと言えるのではないだろうか。

母語別でエラータイプを見た場合、偏りはほぼ想定のとおりに見られた。中国語話者は「ら行」、「な行」、「母音連続」などのエラータイプが多かった。ベトナム語話者は「だ行」、「な行」、「ら行」の子音に関わるエラーが多く、韓国語話者は、拗音化や長音の二重母音化が多く見られた。しかしながら、中国語の三声のような跳ね上がる発音、または、ベトナム語のような鼻にかかるような発音などに対しては、音声認識による分別が困難であるため、コンテンツに盛り込むことができなかった。これらは、教室での運用時の課題として残った。

## 8. 実装結果と課題

「7.3 調査結果」における結果を反映させ、エラー情報を付与した読み上げ文40センテンスをAmiVoiceに実装し、教員がテスト運用を行った。実際に使用する可能性が高い読み上げ文が加わり、また比較的読み上げ文が長くなったことで、中・上級のレベルにあった難易度が確保でき、学習意欲の向上にも良い影響を与えることが期待できた。しかし、同時に、長い読み上げ文には弱点があることが分かった。

長い読み上げ文の場合は、評価の難易度設定を高くしても、図5①において、実際以上に高い評価が出てしまう傾向にある。調べると、この現象が起きる原因は、読み上げ音声に含まれる正しい音素数に対するエラーの音素数の少なさにあることが分かった。中・上級レベルの日本語話者の場合、コミュニケーションに支障が出るほどの発音エラーが出る箇所は、元々あまり多くはない。例えば40拍のセンテンスを読み上げた時、発音の誤りを一ヶ所（一拍）したとする。しかし、残りの39拍が正しければ、母数に対してのエラーの数は1となるため、エラーの出現率は低いものとなる。反対に、短いセンテンスで同じエラーがあった場合は、正しい音素の母数が小さくなるため、結果として、は長いセンテンスよりも評価は低くなる。このような特性を考えると、細かい単音の確認は、短い語彙やセンテンスで行い、追加コンテンツにおいては、モデル音声、入力音声、プロソディグラフ、ピッチ曲線を利用したイントネーションの練習に重きを置

くというように、目的を分けた使用の検討が必要であると考えられる。

また、韻律学習のため、ピッチ曲線を留学生が見るのであれば、ある程度ピッチ曲線を見る際の注意点を指導することも必要である。入力音声のピッチ曲線を見ると、文頭、語頭部分ではピッチ曲線が聴覚印象よりも大きく跳ね上がることが多かった。これは、発話の開始時や語頭では、より大きな力がかかるため、それがピッチ曲線に反映されるためであると思われる。「おそ下がり」と同様、運用の際の際には、この部分についても並行した補足指導が必要である。

自学自習という観点から見ると、本コンテンツは、音声等のクラスにおいて指導を受けた経験のない学生が利用するには、上で見たような学習上、利用上の注意が必要であると言える。継続的な音声学習ができる環境を整えるためには、クラスで指導を受けた留学生が使う場合と、クラスでの学習経験がない留学生が自学自習で利用する場合の2つのケースを視野に入れたマニュアル作成が必要であると考えられた。

## 9. 今後に向けて

本研究では、留学生を取り巻く環境から音声教育の課題を概観し、中・上級レベルの留学生を対象とした音声教材の開発について述べた。音声認識技術を応用した日本語 CALL システムを使うことで、単音の向上や学習意欲の向上が見られたが、中・上級レベルにおける活用の面では、韻律指導の方法、比較的長い文章における練習、自学自習に向けての整備、コースデザインの見直しなどの面において改善しなければならない点が多々あることも明らかになった。分析面では、英語話者やヨーロッパの言語を母語とする留学生の音声の分析、各レベル別のエラー頻度の分析、各音素のエラーの認識率の調査がさらに必要であり、これらは今後の課題であると考ええる。

日本国内においては、一つのクラスに複数の地域からの留学生が在籍していることが多く、母語や出身地域などの背景は様々である。そのため、特に発音のクラスにおいては、学習上の課題として共通する課題もあるが、大きな偏りも存在している。そのような教室環境において、CALL システムの利用は、教室における学習者の心理的負担を軽減し、個々の練習時間の量的確保を可能にする。留学生が長期間、自学自習を継続していく力を養うため、音声のクラスだけではなく、コースデザイン全体をどのように

改善することができるか、さらに考えていく必要がある。

## 参考文献

- アジア人材資金構想プロジェクトサポートセンター (2011) 『教育機関のための外国人留学生就職支援ガイド』 経済産業省
- 阿栄娜・林良子 (2010) 「シャドーイング練習による日本語発音の変化—モンゴル語・中国語母語話者を対象に—」 『電子情報通信学会技術研究報告』 109 (451), 19-24
- 太田由紀子・工藤美佳 (2012) 「中上級の留学生を対象とした日本語音声教育の実践事例—『AmiVoice CALL Web - Japanese-』を使った発音矯正指導—」 『別府大学日本語教育研究 別府大学日本語教育研究センター紀要』 (2), 別府大学日本語教育研究センター, 19-27
- 奥村匡子 (2016) 「日本語学習者に対する音声教育—AmiVoice CALL Web: Japanese を活用した発音学習支援の実践から—」 神奈川大学外国語科目教育部会ワークショップ発表資料
- 小河原義朗 (2001a) 「日本語非母語話者の話す日本語の発音に対する日本人の評価意識: 日本人大学生の場合」 『日本語教育方法研究会誌』 8 (1) : 28-29.
- 小河原義朗 (2001b) 「日本語非母語話者の話す日本語の発音に対する日本人の評価意識: 社会人の場合」 『日本語教育方法研究会誌』 8 (2) : 10-11.
- 小河原 (2009) 「多様化する日本語教育における音声教育の目標と教師の役割をとらえ直す」 河野俊之, 小河原義朗編 『日本語教育の過去・現在・未来 第4巻 音声』 凡人社, 48-69
- 河野俊之 (2015) 「自己モニターを活用した音声教育における独自の基準の観点」 『日本語教育方法研究会誌』 vol.22 No.1, 2015, 100-101
- 河野俊之 (2016) 「単純名詞のアクセント教育の実践」 『日本語教育方法研究会誌』 vol.23 No.1, 2016, 14-15
- 河野俊之・串田真知子・築地伸美・松崎寛 (2004) 『1日10分の発音練習』 くろしお出版
- 佐藤友則 (1994) 「単音と韻律が日本語音声の評価に与える影響力の比較」 『世界の日本語教育』 Vol.5, 139-154
- 高橋恵利子・松崎寛 (2007) 「プロソディ・シャドーイングが日本語学習者の発音に与える影響」 『広島大学日本語教育研究』 17, 73-80
- 萩原廣 (2005) 「日本語の発音指導におけるシャドーイングの有効性」 『京都経済短期大学論集』 13 (1), 55-71
- 古井貞熙 (2000) 「音声認識技術の現状と課題」 『音声研究』 第4巻, 第3号, 60-63
- 松崎寛・高橋恵利子 (2015) 「お下がりに対応したプロソディグラフ自動出力システムの開発」 『日本語教育方法研究会誌』 Vol.22, No.1, 48-49
- 牧野成一・鎌田修・山内博之・齋藤真理子・萩原稚佳子・伊藤とく美・池崎美代子・中島和子 (2001) 『ACTFLOPI 入門—日本語学習者の『話す』を客観的に測る』 アルク,



**参照 URL**

アドバンスド・メディア「音声認識とは」

<http://www.advanced-media.co.jp/amivoice> (2016年9月1日参照)

岸慶騎・中村雅巳(株式会社アドバンスド・メディア)(2002)「AmiVoice Medical Solution ～医療現場での音声認識活用～」Seagaia Meeting2002

<http://www.seagaia.org/sg2002/HTML/kishi/kishi.html>

(2016年9月19日参照)

日本学生支援機構「外国人留学生進路状況・学位授与状況調査」平成22年から平成26年

[http://www.jasso.go.jp/about/statistics/intl\\_student\\_d/index.html](http://www.jasso.go.jp/about/statistics/intl_student_d/index.html)

(2016年9月10日参照)

**参考資料**

株式会社アドバンスド・メディア(2011)「AmiVoice CALL Web – Japanese-v1.0 操作マニュアル(学習者)」

**註**

- i 牧野他(2001:19)表2-②正確さ,表2-③正確さによる
- ii 牧野他(2001:19)表2-②正確さ,表2-③正確さによる
- iii 図1,図2は,日本学生支援機構「外国人留学生進路状況・学位授与状況調査」平成22年～平成26年の数値をまとめたもの。進路先には「日本で就職」,「日本で進学」,「日本国内その他」,「出身国(地域)で就職」,「出身国(地域)で進学」,「出身国(地域)その他」,「日本・出身国(地域)以外で就職」,「日本・出身国(地域)以外で進学」,「日本・出身国(地域)以外その他」の分類となっている。本稿では,大学「専修学校」「準備教育過程」
- iv 図1と同様。「その他」の進路先は除外した。
- v アドバンスド・メディア「音声認識とは」(2016年9月1日参照)  
<http://www.advanced-media.co.jp/amivoice>
- vi 音声出力までの過程については,岸・中村(2002)を参照のこと。
- vii AmiVoice CALL Web 操作マニュアル p.18
- viii 河野・串田・築地・松崎(2004)で用いられるようになった韻律表示の方法。ピッチ曲線から,音節単位で区切り,韻律情報を付加して表示したもの。
- ix 2018年3月までの使用が可能。  
<http://www.u.tsukuba.ac.jp/~matsuzaki.hiroshi.jp/pg.html>
- x 授業で発表する場面,大学の事務での問い合わせの場面,先生にお願いする場面,図書館での場面,アルバイト・就職活動における場面,会社での挨拶の場面を設定した。
- xi 巻末にあげた資料は「AmiVoice CALL Web -Japanese- v1.0」の操作マニュアル(学習者)であるが,エラータイプは34である。実際に使用したのは「AmiVoice CALL

Web-Japanese-v1.1」であり、エラータイプは38である。

- xii 留学生の母語は、中国語20名、ベトナム語6名、韓国語3名、ネパール語2名であった
- xiii 発音の誤りが起こる可能性は、一カ所につき一つのエラーというものではなく、複数のエラータイプが起こる可能性がある。例えば「資料」は、「しろー」、「しりよー」、「しーりょ」、「しりょ」のように発話者によって異なるエラーが出る。そこで、ここでは、エラー箇所の数ではなく、実際に起こったエラーの数を数えた。

巻末資料

発音評価が可能な 34 エラータイプ (AmiVoice CALL Web -Japanese- v1.0 操作マニュアル 学習者)

エラータイプ	説明
○長い音→×短い音	長い音が短い音になっています。
○短い音→×長い音	短い音が長い音になっています。
「っ」:○ある→×ない	「っ」がありません。
「っ」:○ない→×ある	「っ」があります。
○有声音→×無声音	有声音(「 <sup>h</sup> 」がある音)が無声音(「 <sup>h</sup> 」がない音)になっています。
○無声音→×有声音	無声音(「 <sup>h</sup> 」がない音)が有声音(「 <sup>h</sup> 」がある音)になっています。
○「つ」→×「ちゅ」	「つ」が「ちゅ」になっています。
○「つ」→×「す」	「つ」が「す」になっています。
母音:○ない→×ある	母音 (i, u) を発音しています。
○ら行→×な行	ら行 (らりるれろ) が, な行 (なにぬねの) になっています。
「ん」:○ある→×ない	「ん」がありません。
「ん」:○ない→×ある	「ん」があります。
○「し(じ)」→×「ち」	「し(じ)」が「ち」になっています。
○「ち」→×「し(じ)」	「ち」が「し(じ)」になってます。
○だ行→×ら行	だ行 (だでど) が, ら行 (られろ) になっています。
○は行→×あ行	は行 (はひふへほ) が, あ行 (あいうえお) になっています。
「ん」+ a	「ん」と次の音がいっしょになって, 他の音になっています。
○直音→×拗音	直音 (「やゆよ」がない音) が拗音 (「やゆよ」がある音) になっています。
母音連続	母音 (a, i, u, e, o) が続く時, 他の音になったり母音がなくなったりしています。
○「ん」/「っ」→×長い音	「ん」や「っ」が長い音になっています。
○長い音/「ん」→×「っ」	長い音や「ん」が「っ」になっています。
○「っ」/長い音→×「ん」	「っ」や長い音が「ん」になっています。
○長い音→×二つの母音	長い音が二つの母音 (a, i, u, e, o) になっています。
○う段→×お段	う段 (う, く, す, つ・・・) が, お段 (お, こ, そ, と・・・) になっています。
○お段→×う段	お段 (お, こ, そ, と・・・) が, う段 (う, く, す, つ・・・) になっています。
○拗音→×直音	拗音 (「やゆよ」がある音) が直音 (「やゆよ」がない音) になっています。
○や行→×じゃ行	や行 (やゆよ) が, じゃ行 (じゃじゅじょ) になっています。
○や行→×ざ行	や行 (やゆよ) が, ざ行 (ざずぞ) になっています。
○い段→×え段	い段 (い, き, し, ち・・・) が, え段 (え, け, せ, て・・・) になっています。
○え段→×い段	え段 (え, け, せ, て・・・) が, い段 (い, き, し, ち・・・) になっています。
○な行→×ら行	な行 (なにぬねの) が, ら行 (らりるれろ) になっています。
○か行→×は行	か行 (かきくけこ) が, は行 (はひふへほ) になっています。
○「やゆよ」→×「やゆよ」	「やゆよ」が「やゆよ」になっています。
○た行→× [ts]	た行 (たてと) のはじめが [ts] のように聞こえます。